## AMENDMENTS TO THE CLAIMS

**This listing of claims will replace all prior versions and listings of claims in the application:**

## LISTING OF CLAIMS:

1. - 19. ( cancelled)

20. (currently amended): A method for identifying a candidate protein useful as an anti-infective, comprising:

(a) calculating computationally protein sequence-based attributes from protein sequences of a pathogenic organism, wherein said protein sequences are predicted either from whole ~~genomic sequences,~~ or ~~from~~ partial genomic sequences ~~comprising at least one chromosome~~, and wherein said protein sequence-based attributes comprise: percentage of charged amino acids, percentage hydrophobicity, distance of protein sequence from a fixed reference frame, measure of dipeptide complexity, and measure of hydrophobicity from a fixed reference frame, and wherein said pathogenic organism is selected from the group consisting of *B.burgdorfei*, *C.jejuni*, *C.pneumoniae*, *C.trachomatis*, *H.influenzae*, *H.pylori*, *L.major*, *M.genitalium*, *M.pneumoniae*, *M.tuberculosis*, ~~*N.menigitis*~~ *N.meningitidis*, *P.aeruginosa*, *P.falciparum*, *R.prowazekii*, *T.pallidum*, and *V.cholerae*;

(b) clustering computationally said protein sequences based on said protein sequence-based attributes using Principle Component Analysis;

(c) identifying computationally outlier ~~proteins~~ protein sequences, wherein said outlier ~~proteins~~ protein sequences appear outside a main cluster;

(d) comparing said outlier protein sequences to protein sequences listed in public sequence databases of ~~a group of pathogenic organisms consisting~~all organisms including of *B.burgdorfei*, *C.jejuni*, *C.pneumoniae*, *C.trachomatis*, *H.influenzae*, *H.pylori*, *L.major*, *M.genitalium*, *M.pneumoniae*, *M.tuberculosis*, ~~*N.menigitis*~~ *N.meningitidis*, *P.aeruginosa*, *P.falciparum*, *R.prowazekii*, *T.pallidum*, and *V.cholerae* to (1) identify outlier proteins that are unique to said pathogenic organism based on the sequences in the databases accessed for the

comparing, and ~~to~~ (2) identify outlier proteins that are homologous or identical to proteins known to be involved in virulence; and

(e) displaying the results of said step (d).

21. (Canceled).

22. (previously presented): The method of claim 20, wherein said protein sequence based attributes comprise fixed protein attributes and variable protein attributes.

23. (previously presented): The method of claim 22, wherein a variable protein attribute is a distance of protein sequence from a variable reference frame.

24. (previously presented): The method of claim 20, wherein said clustering is done by Principle Component Analysis using correlation coefficient between said protein sequence-based attributes.

25. (Canceled)

26. (currently amended): The method of claim 20, wherein the outlier protein sequences identified in step (d) is non-homologous to known anti-infective proteins from a pathogen selected from the group consisting of *B.burgdorfei*, *C.jejuni*, *C.pneumoniae*, *C.trachomatis*, *H.influenzae*, *H.pylori*, *L.major*, *M.genitalium*, *M.pneumoniae*, *M.tuberculosis*, ~~*N.menigitis*~~ *N.meningitidis*, *P.aeruginosa*, *P.falciparum*, *R.prowazekii*, *T.pallidum*, and *V.cholerae*.

27. (currently amended): The method of claim 20, wherein the outlier protein sequences identified in step (d) has an amino acid sequence selected from the group consisting of SEQ ID Nos: 1-31.

28. (currently amended): The method of claim 20, wherein the outlier protein sequences ~~selected in step (c)~~ identified in step (d) has an amino acid sequence selected from the group consisting of SEQ ID Nos: 32-118.

29. (currently amended): The method of claim 20, wherein steps (a)-(c) are performed by a computer system ~~comprising:~~ that

~~(1) a central processing unit (CPU), wherein said CPU~~ executes a program that calculates protein sequence-based attributes, wherein said protein sequence-based attributes comprise: percentage of charged amino acids, percentage hydrophobicity, distance of protein sequence from a fixed reference frame, measure of dipeptide complexity, and measure of hydrophobicity from a fixed reference frame; and clusters protein sequences based on said protein sequence-based attributes using Principle Component Analysis, thereby producing results~~;~~

~~(2) a memory device accessed by said CPU, wherein said memory device stores said results;~~

~~(3) a display on which said CPU displays said results in response to user inputs; and~~

~~(4) a user interface device~~.

30. - 33. (Canceled).